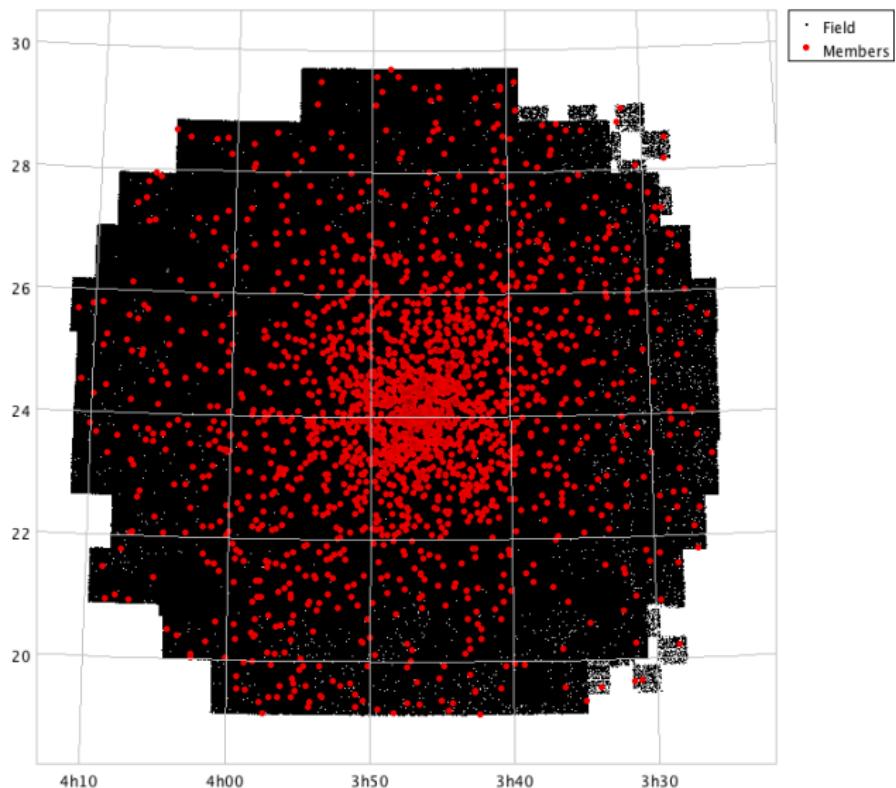


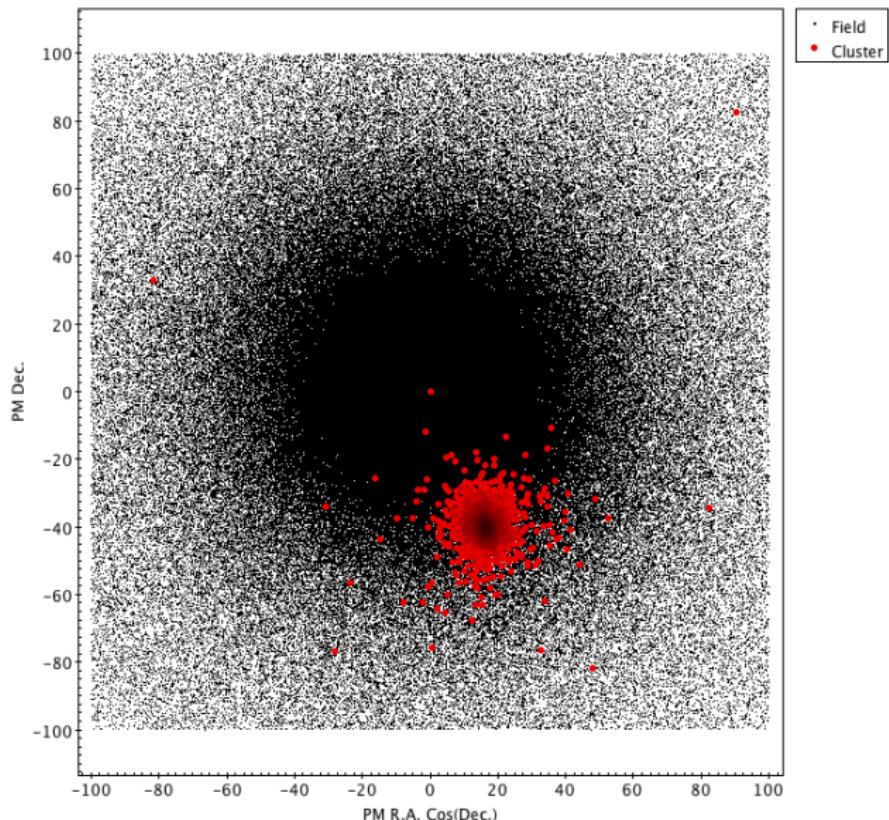
Star clusters



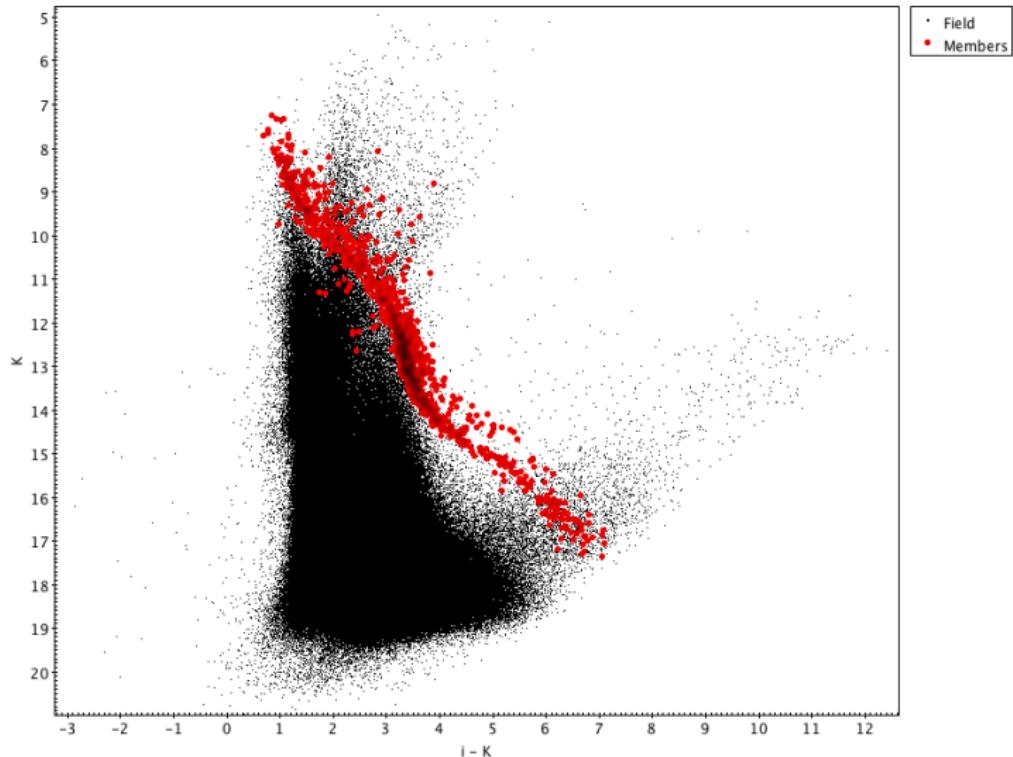
Star clusters



Star clusters



Star clusters



Ideal vs. Reality

- Perfect tags
- Properties of stars
 - 6D phase space (3D positions, 3D velocities)
 - luminosities
 - Metallicites
 - Ages
 - Temperatures and masses
 - Magnetic fields
- All stars
- Precision and accuracy
- Membership probabilities
- Properties of some stars
 - 2D positions, 2D proper motions, parallaxes, radial velocities
 - Apparent magnitudes
 - Metallicities
 - Masses (binaries)
 - ...
- Global properties
- Precision depends on distance/brithness
- Dependent on stellar models



How things are done now?



UNED

- BANYAN I (Malo et al. 2013)
- UPMASK (Krone-Martins & Motinho 2014)
- Sarro (Sarro et al. 2014)
- BANYAN II (Gagne et al. 2015)
- Sampedro (Sampedro and Alfaro 2016)
- LAceWing (Riedel et al. 2017)

Approaching the ideal

- Global properties
- Use as many objects as possible
- Propagate uncertainties

Approaching the ideal

- Global properties
 - Use as many objects as possible
 - Propagate uncertainties
 - Parametric distributions
 - Modelling missing values
 - Individual uncertainties
-
- 1 Big data
 - 2 All-Inclusive

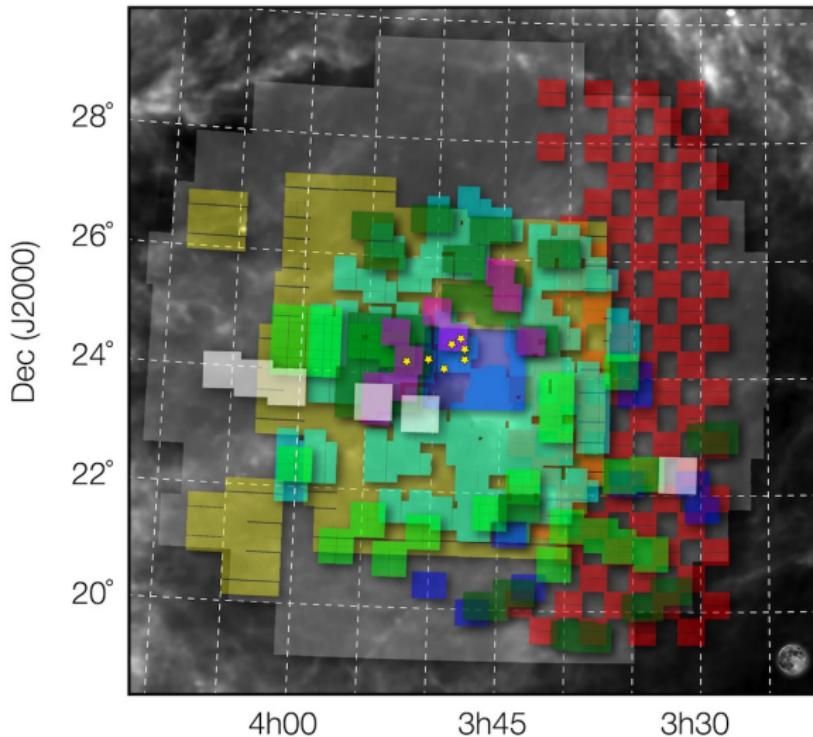
- 1** Choose a Cluster: the Pleiades.
- 2** Select a data set: DANCe DR2 (DDR2).
- 3** Create the methodology
- 4** Validate and get results
- 5** Compare to previous results

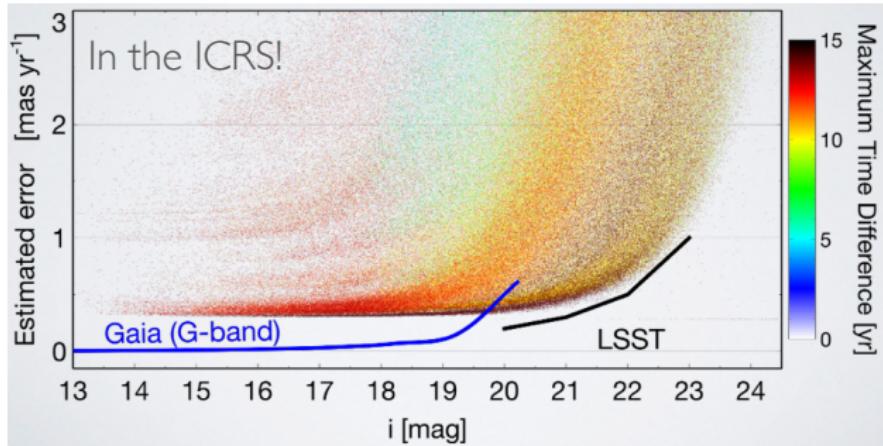
Pleiades History

Sorry ... Next time!

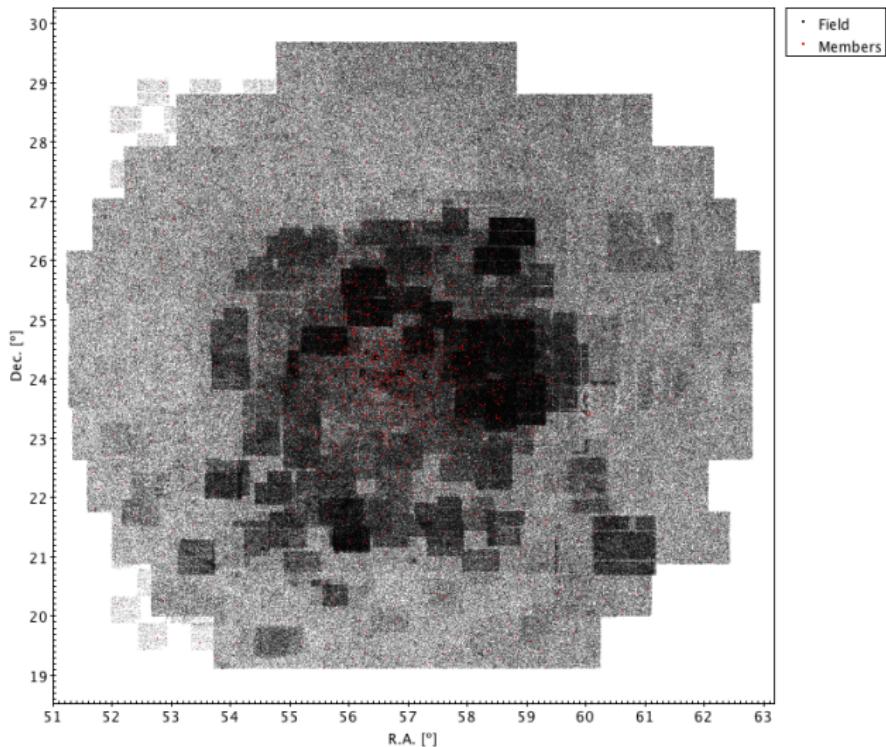
DDR2(Bouy+2013)

KPNO/Mosaic1 UKIRT/WFCAM Subaru/SuprimeCam CFHT/CFHT12K
INT/WFC CFHT/UH8K KPNO/NEWFIRM CTIO/MOSAIC2 CFHT/MegaCam

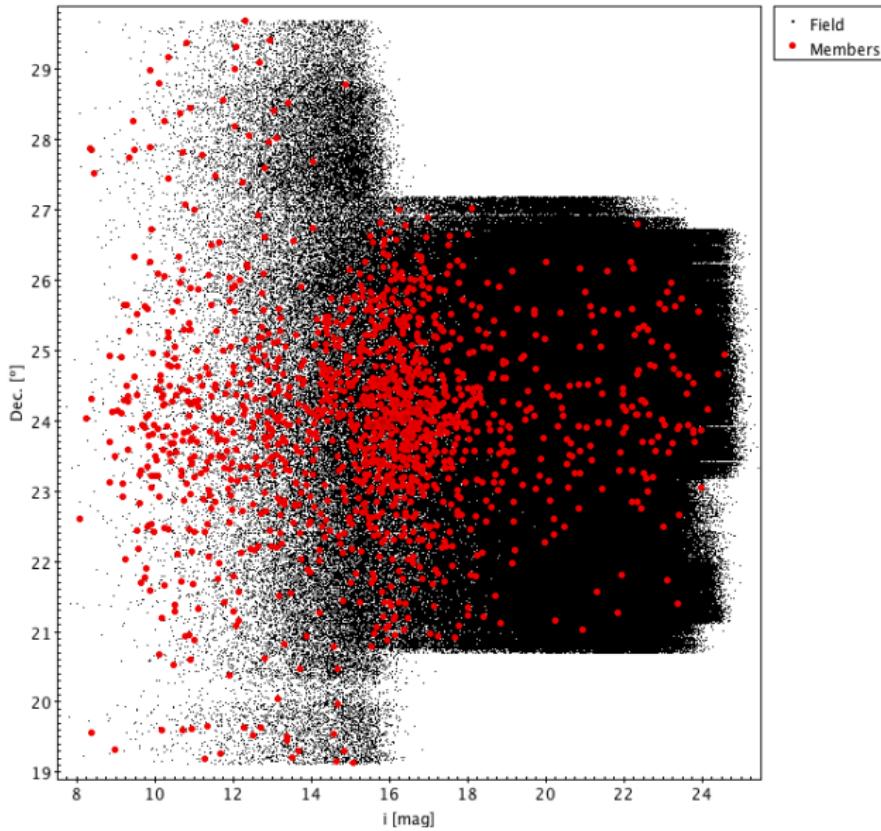




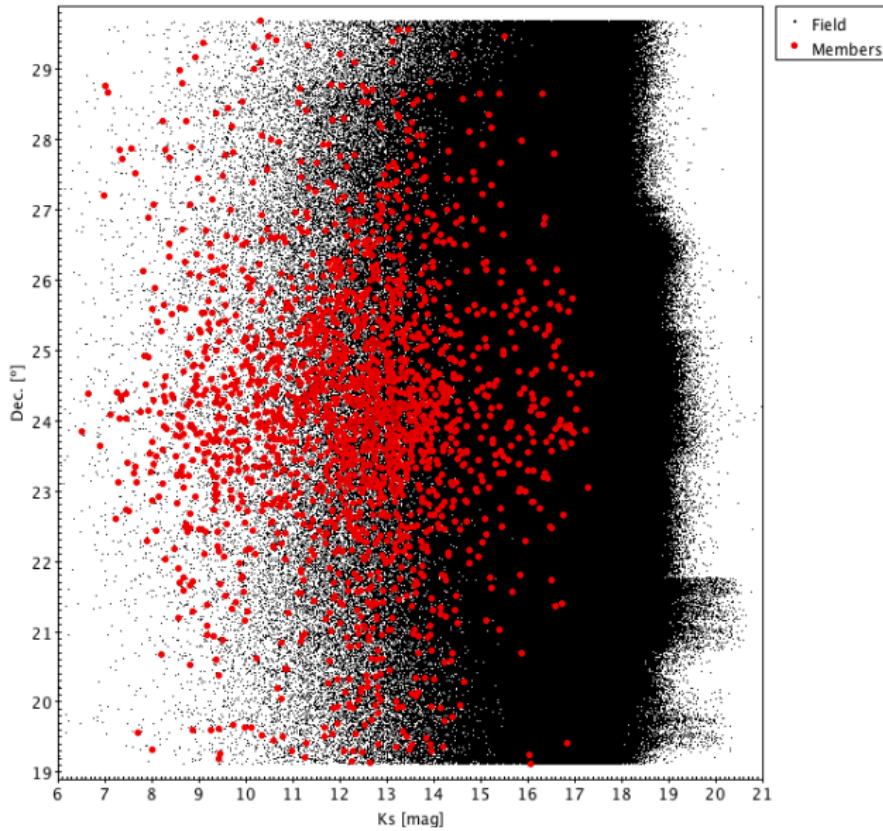
DDR2(Bouy+2013)

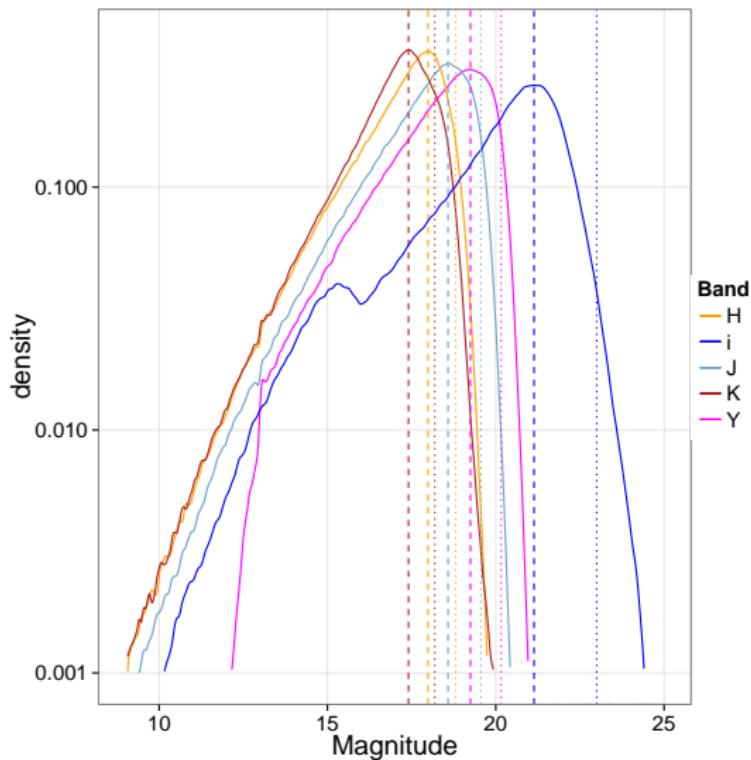


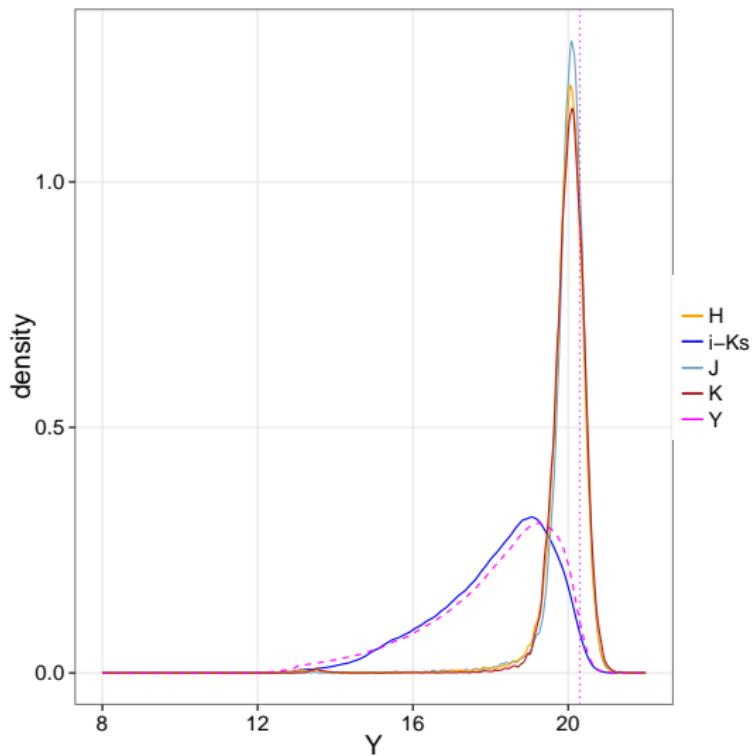
DDR2(Bouy+2013)

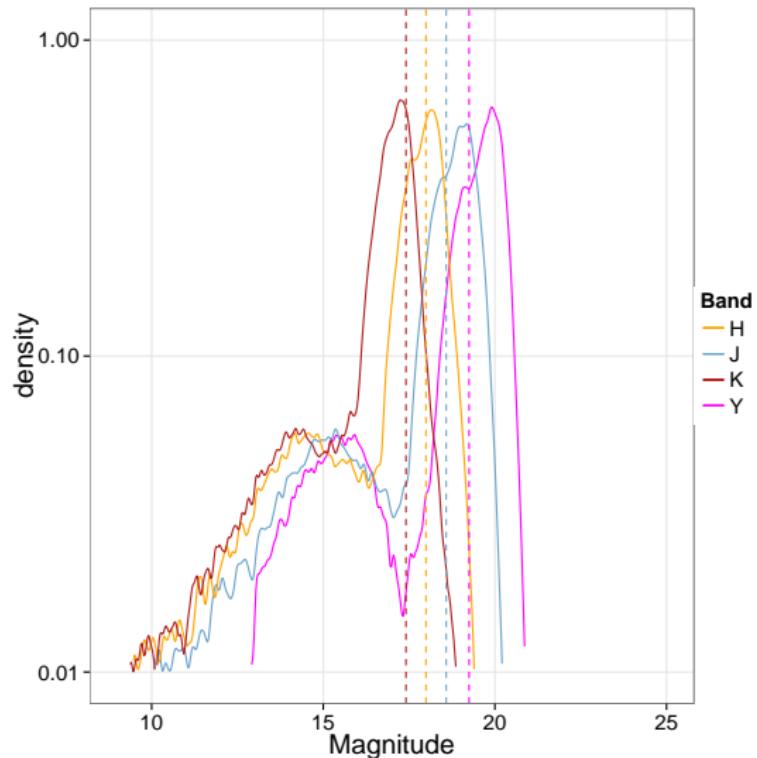


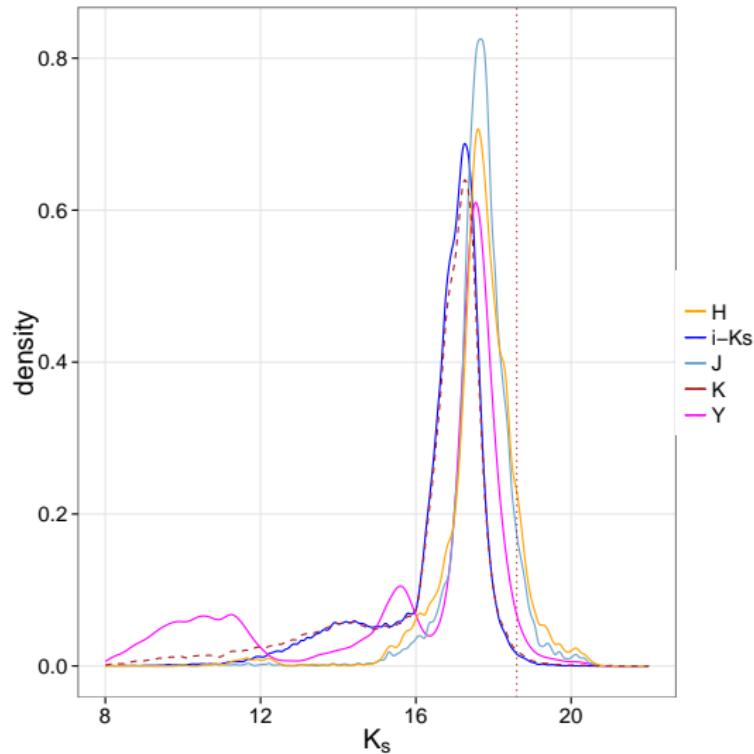
DDR2(Bouy+2013)

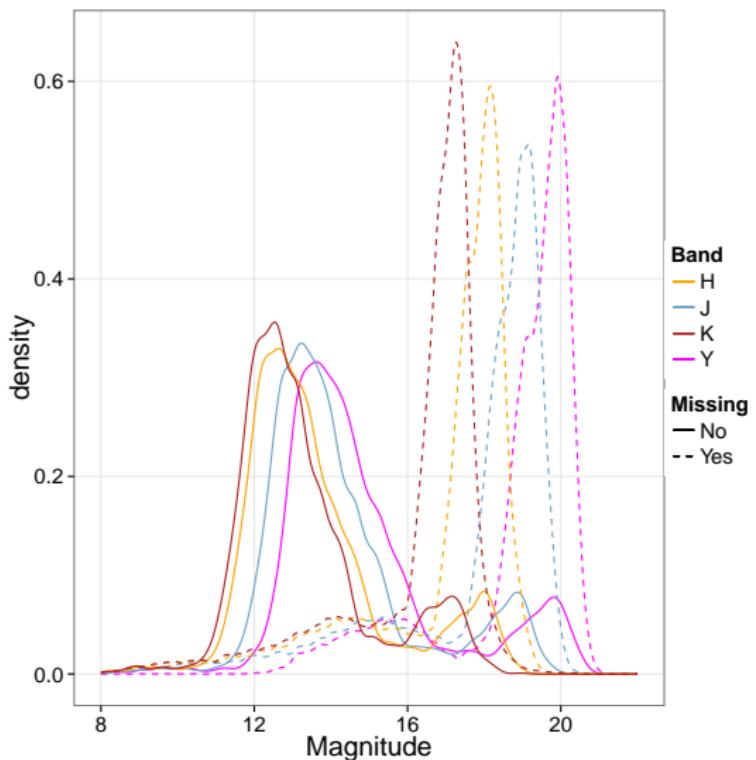












Selection of observables:

- Sarro et al. 2014 -> Proper motions and $i - K_s, Y - J, J, H, K_s$
 - Higher exhaustivity
 - Lower contamination
 - Conciseness
- Analysis of best models for spatial distribution

Generative model

$$p(\mathbf{d}_n | \pi, \boldsymbol{\theta}_c, \boldsymbol{\theta}_f, \mathbf{u}_n) = \pi \cdot p_f(\mathbf{d}_n | \boldsymbol{\theta}_f, \mathbf{u}_n) + (1 - \pi) \cdot p_c(\mathbf{d}_n | \boldsymbol{\theta}_c, \mathbf{u}_n),$$

Membership probabilities

$$p(\mathcal{M}_c | \mathbf{d}_n) = \frac{p_c(\mathbf{d}_n | \boldsymbol{\theta}_c, \mathbf{u}_n) \cdot (1 - \pi)}{\pi \cdot p_f(\mathbf{d}_n | \boldsymbol{\theta}_f, \mathbf{u}_n) + (1 - \pi) \cdot p_c(\mathbf{d}_n | \boldsymbol{\theta}_c, \mathbf{u}_n)},$$

- Photometry independent of proper motions.
- Cluster -> Single stars and Equal-Mass Binaries (EMB)

Types:

- Deterministic
- Stochastic

Flavours:

- Probability of detection
- Probability of selection

Simplistic assumptions:

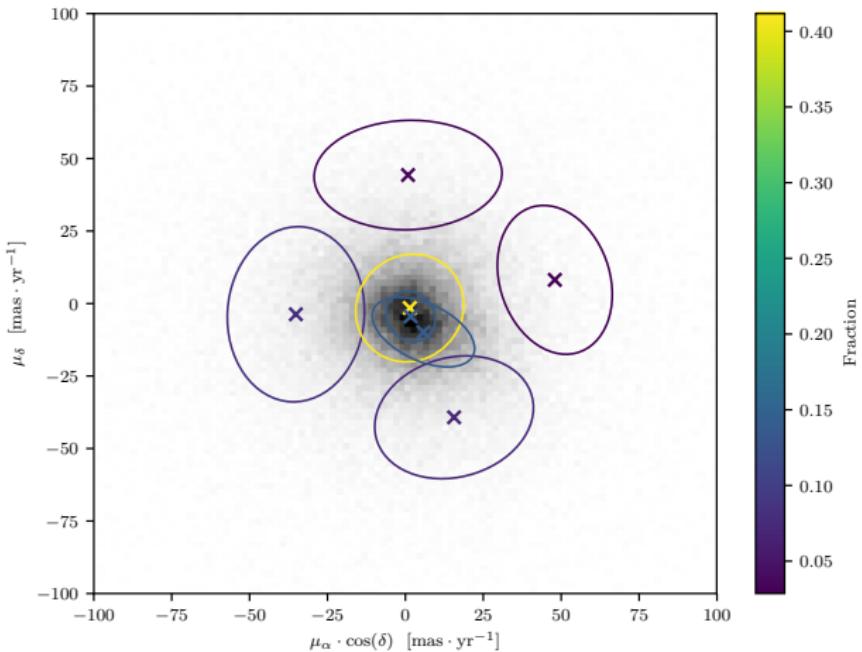
- Field => Ignorability
- Cluster => Missing at random

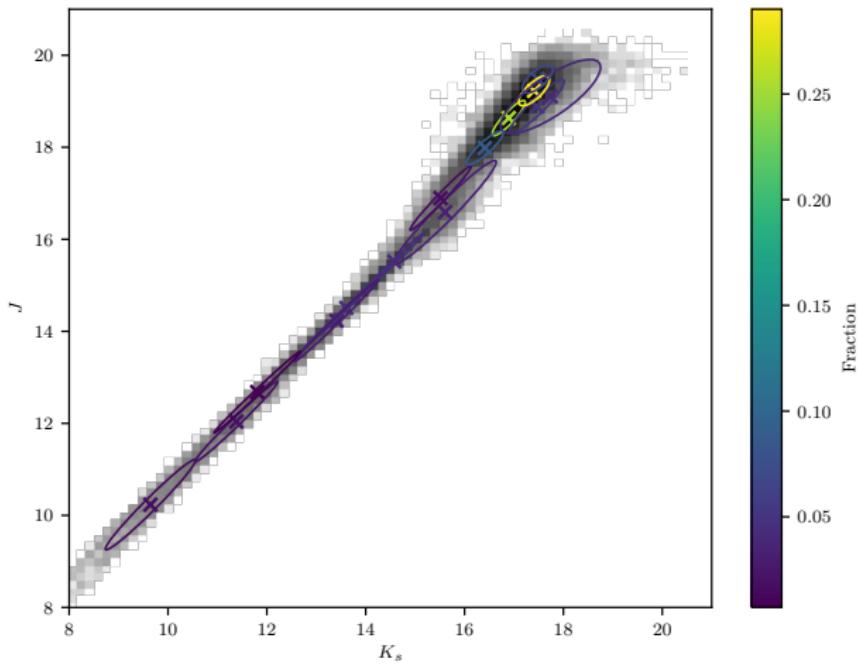
Model:

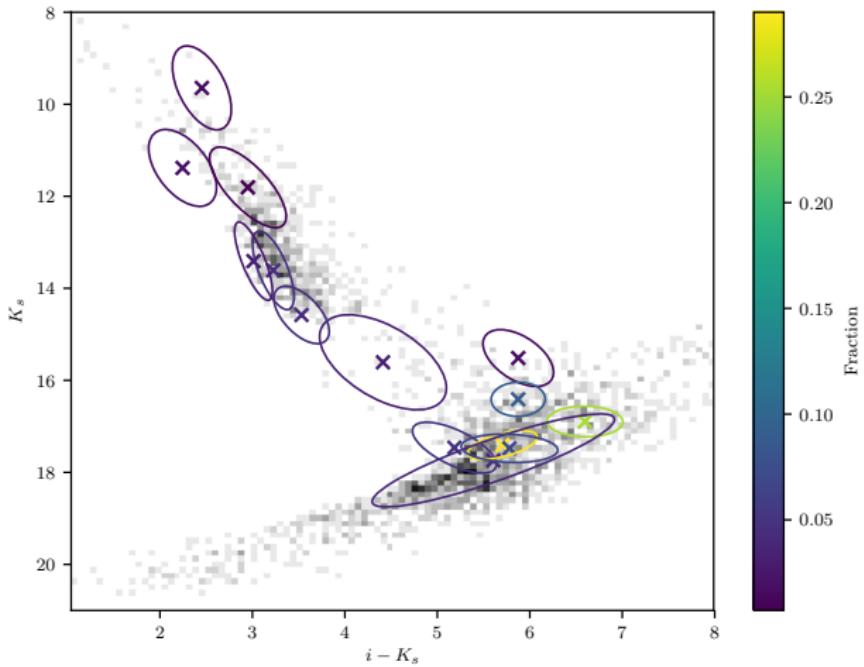
$$p_{GMM}(x|\pi, \mu, \Sigma) = \sum_{m=1}^M \pi_m \cdot \mathcal{N}(\mu_m, \Sigma_m),$$

Fitting GMMs:

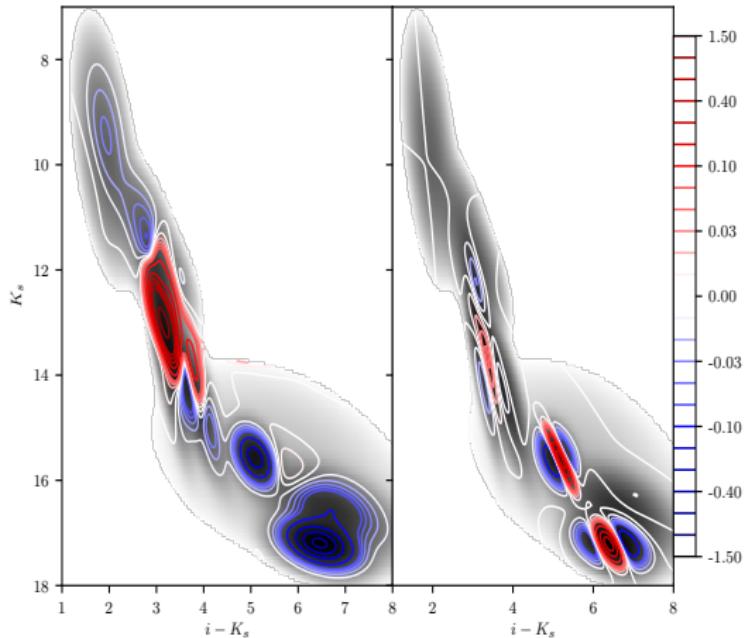
- Proper motions => MLE
- Photometry => MLE with missing values







Ignorability vs Naivety



- RMSRD density
 - Naivety: 0.78 ± 0.38
 - Ignorability: 0.21 ± 0.4
- MAR parameters
 - Naivety: 1.5 ± 3.2
 - Ignorability: 0.42 ± 2.5

Overview

Before:

- One population
- Proper motions (1 or 2)
- Cluster sequence
 - Stellar models
 - PCA
 - Principal Curve (λ)

Now:

- Single stars and EMB
- Proper motions (4 and 2)
- Sequences:
 - Empirical splines
 - Intrinsic (5D) dispersion
 - True colour distribution.



Cluster photometric model

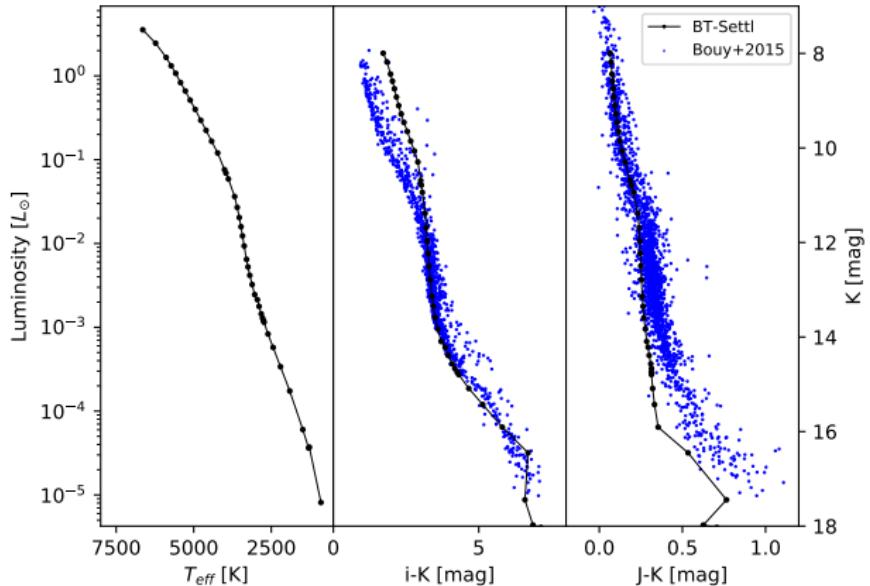


UNED

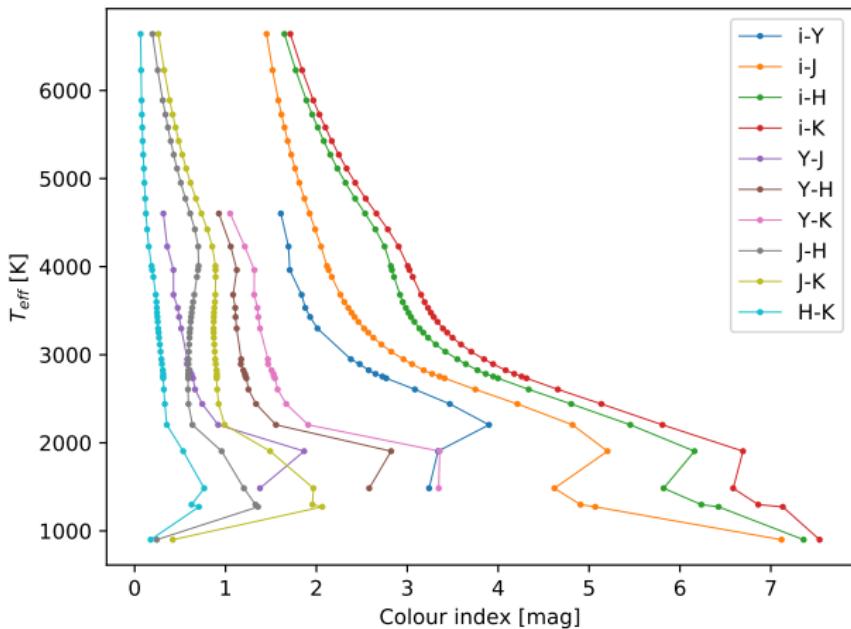
Colour index $i - K_s$:

- 1** Sarro et al. 2014
- 2** Bessel et al. 1998, $T_{eff} <= > M$ dwarfs
- 3** Allard et al. 2014, BT- Settl model

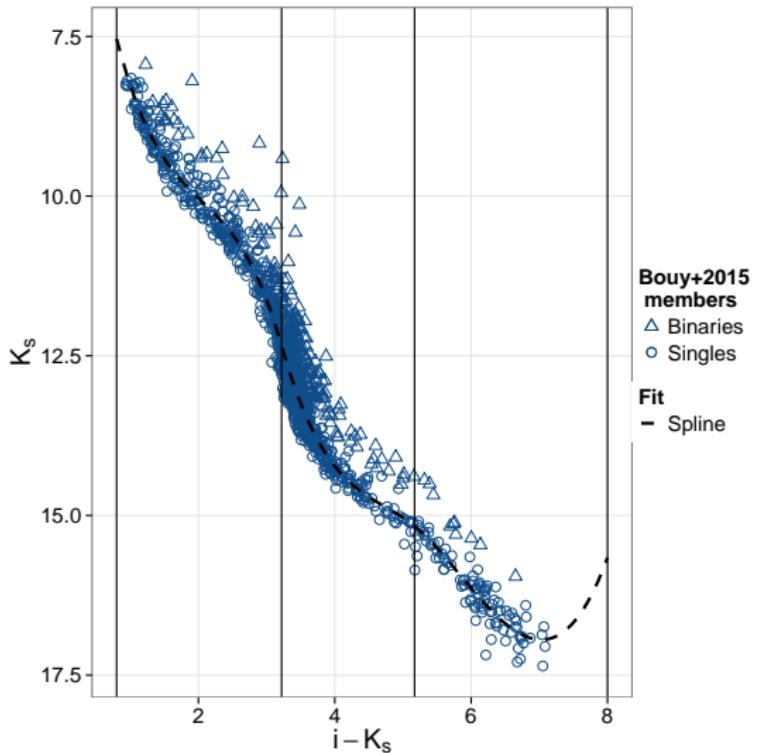
Cluster photometric model



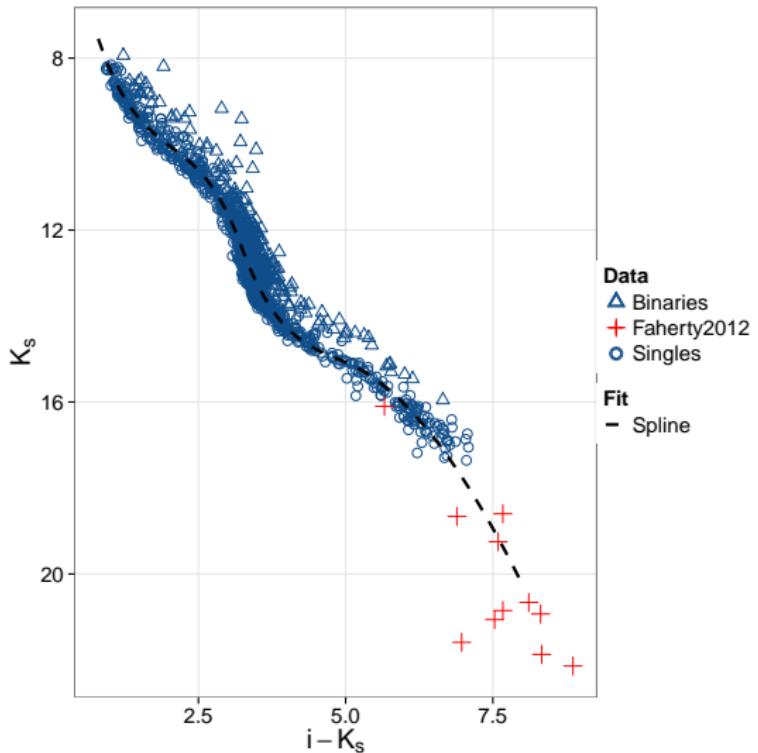
Cluster photometric model



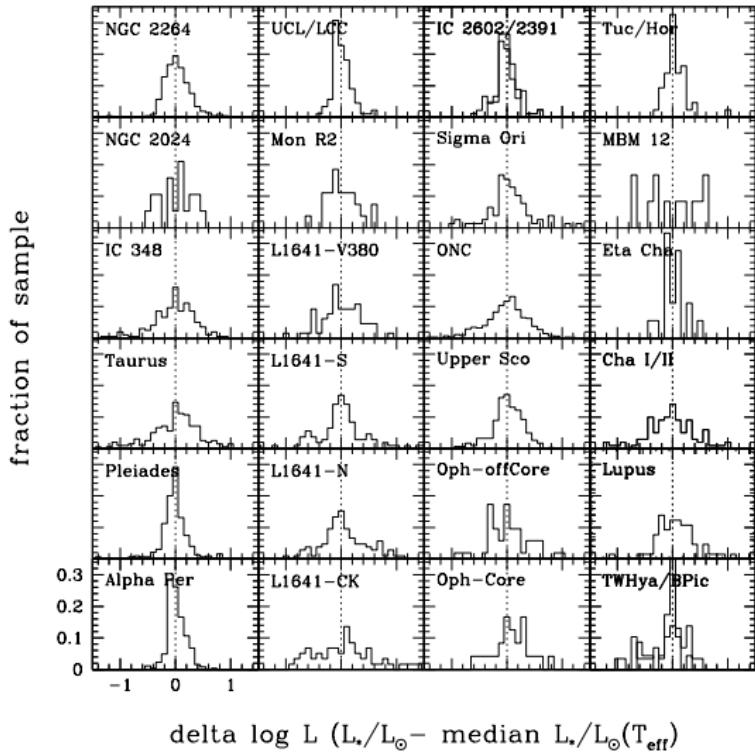
Cluster photometric model



Cluster photometric model



Cluster photometric model



Validation on synthetic data

Five synthetic data sets: with and without missing values.

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

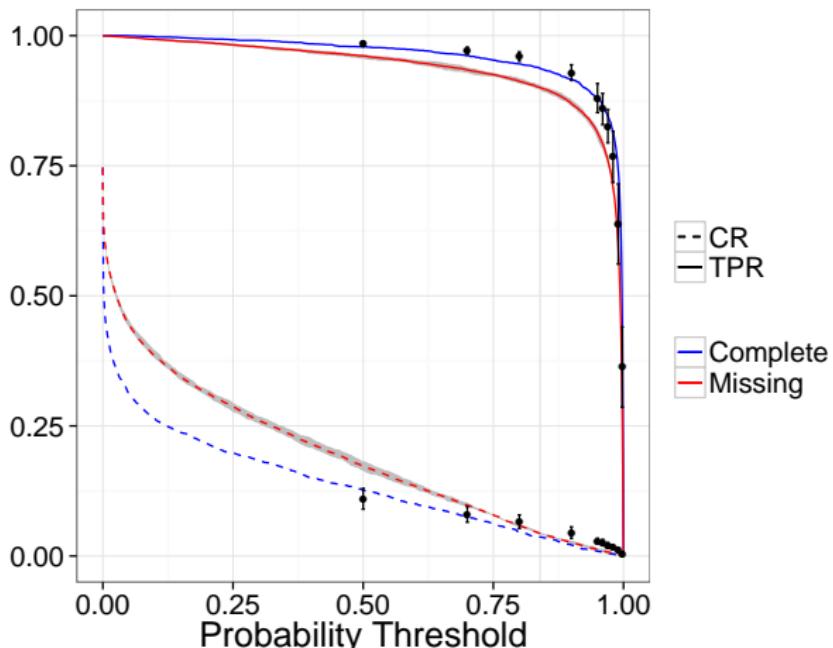
$$CR = \frac{FP}{FP + TP}$$

$$PPV = \frac{TP}{TP + FP}$$

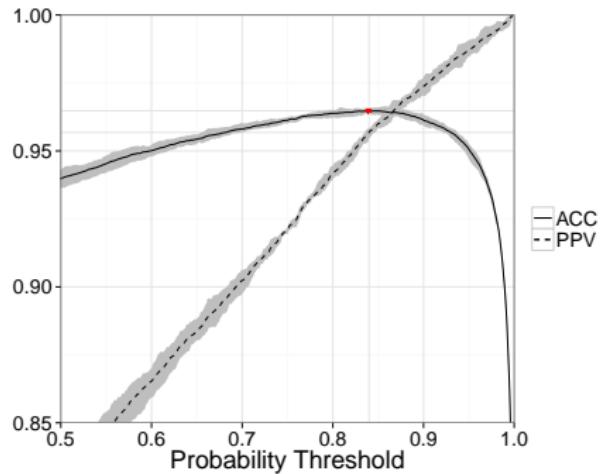
$$ACC = \frac{TP + TN}{TN + FN + TP + FP},$$

Validation on synthetic data

$$\langle CR \rangle = 5.8 \pm 0.2\%$$

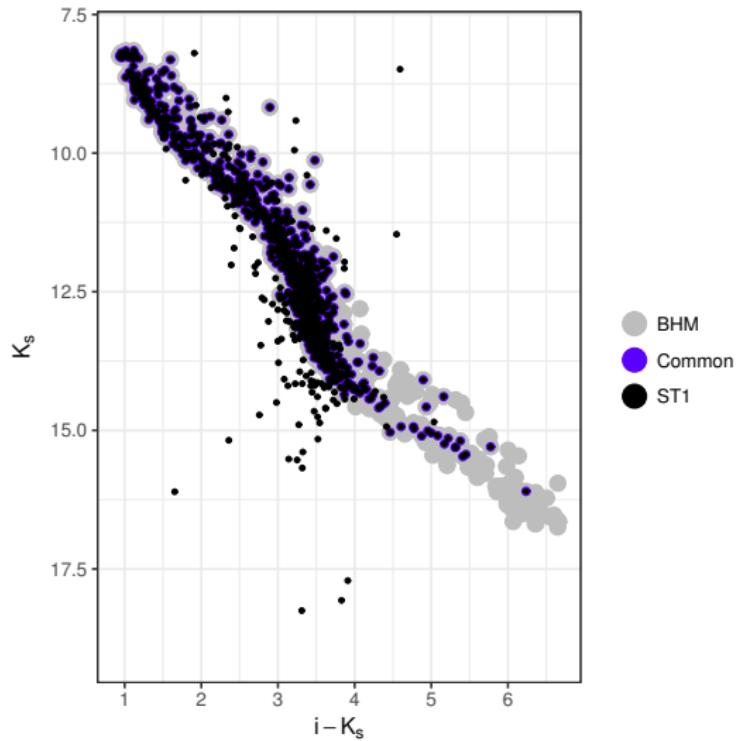


Validation on synthetic data

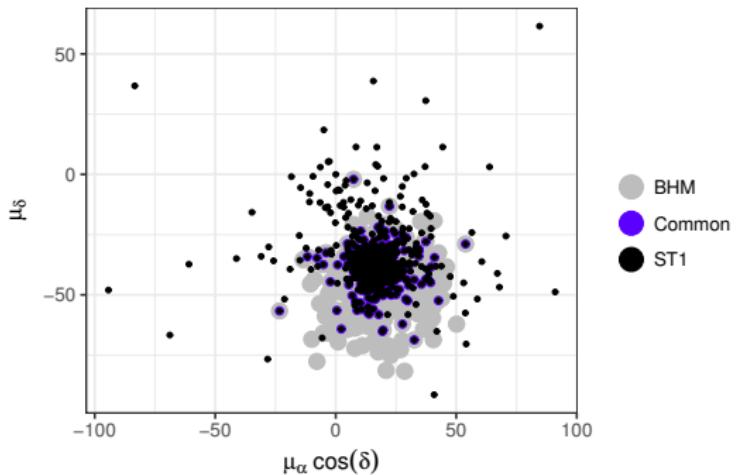


$\text{ACC}=96.5 \pm 0.1\%$, $\text{CR}=4.3 \pm 0.2\%$, $\text{TPR}=90.0 \pm 0.05\%$,
 $\text{PPV}=95.6 \pm 0.2\%$.

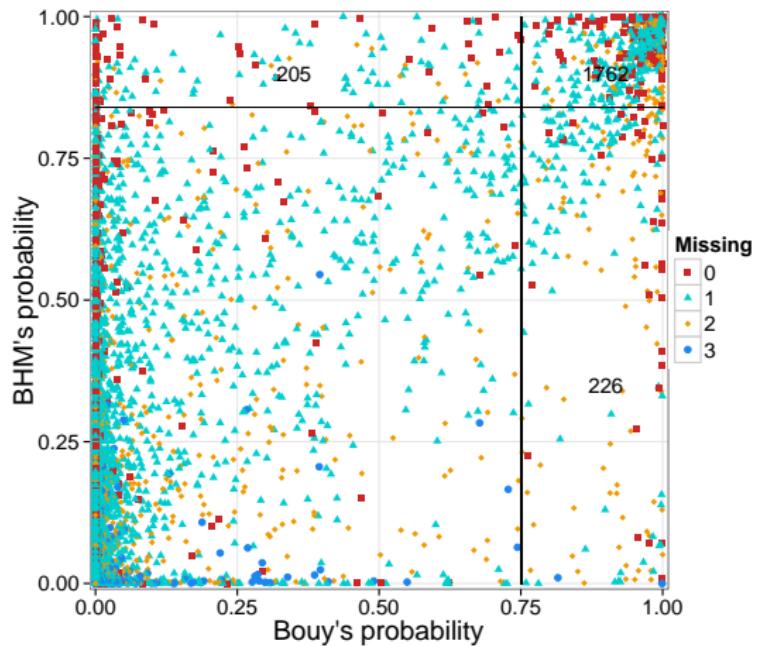
Comparison with literature



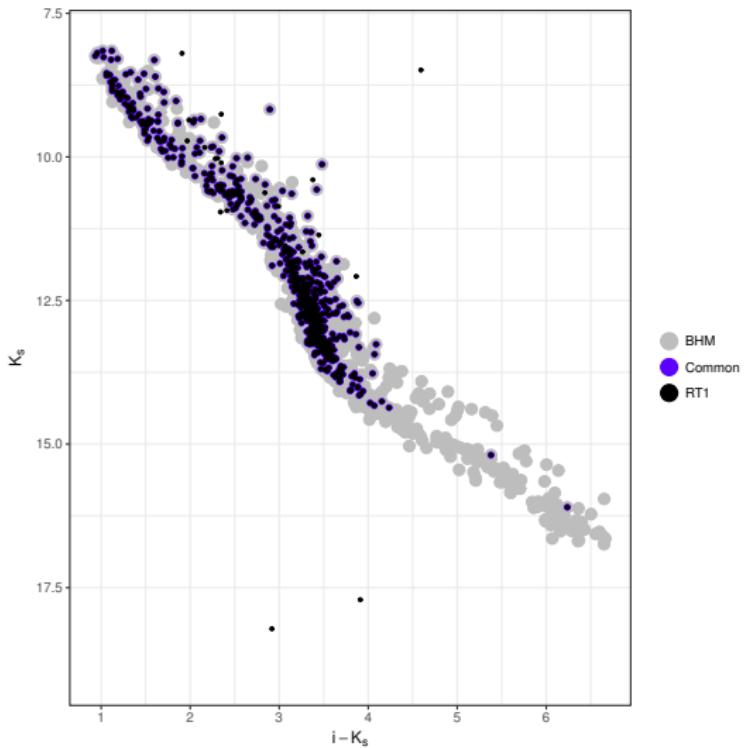
Comparison with literature



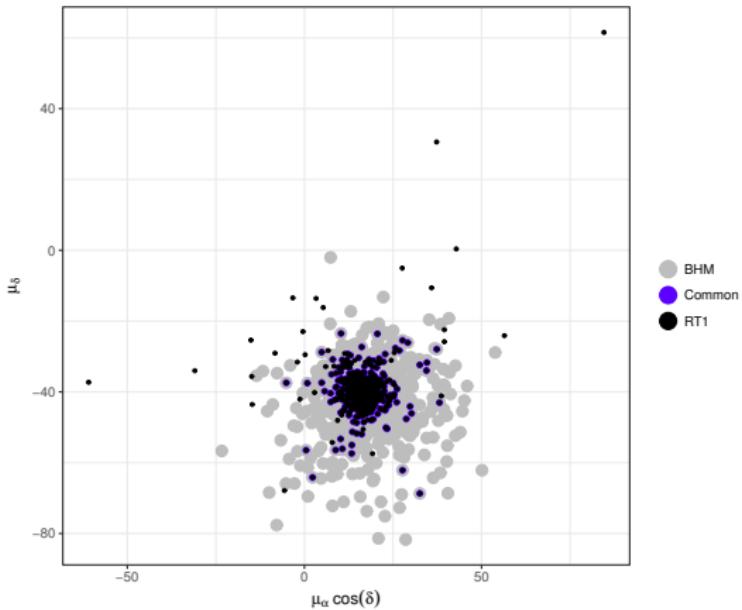
Comparison with literature



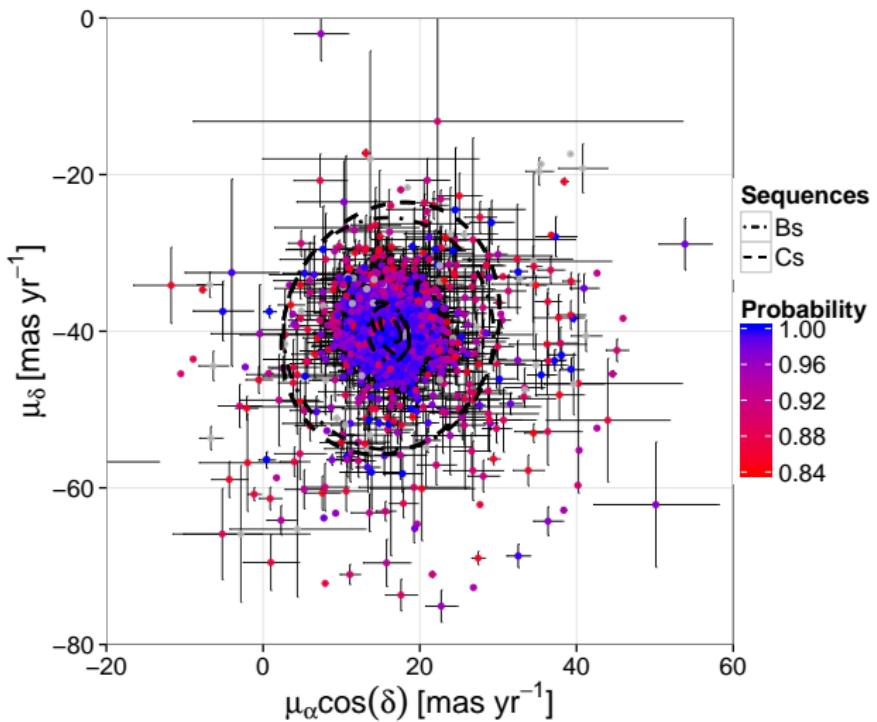
Comparison with literature



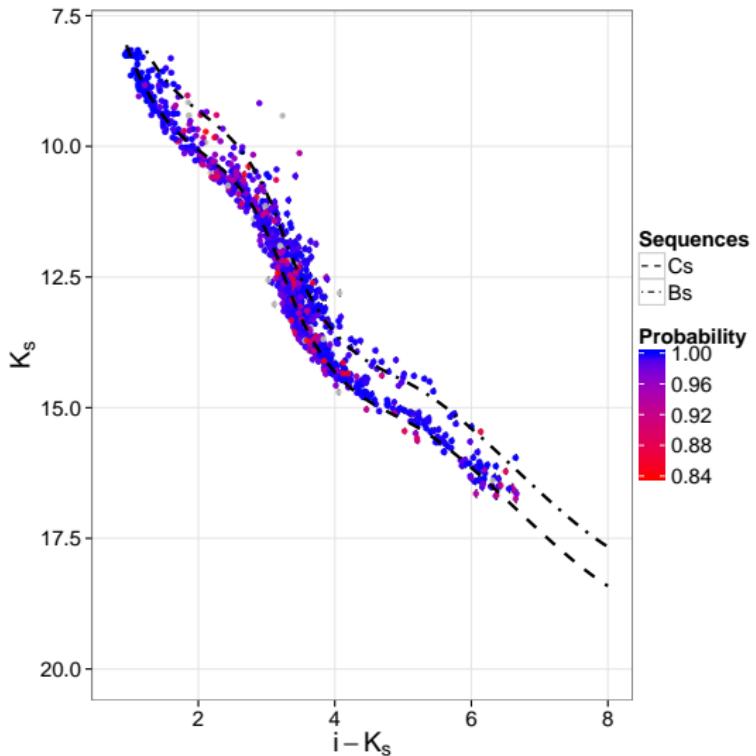
Comparison with literature



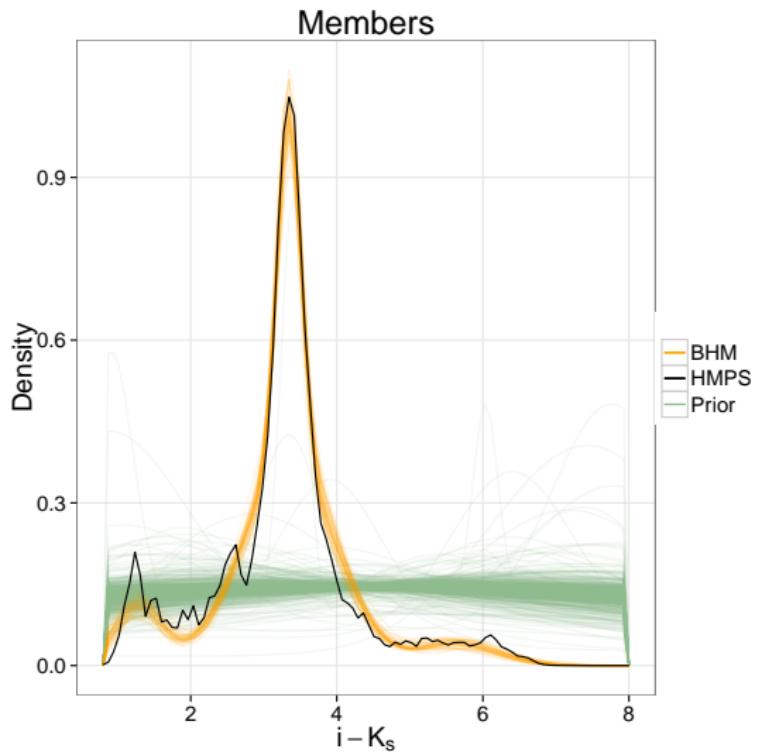
Results



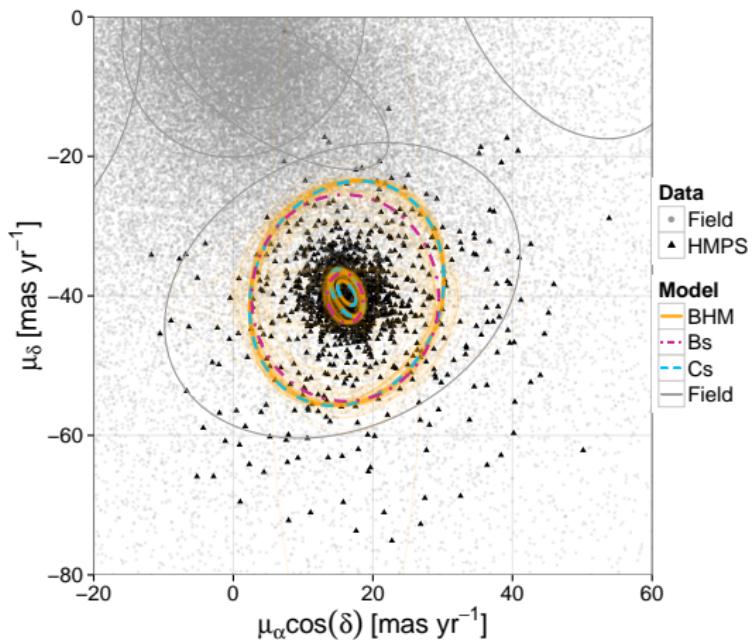
Results



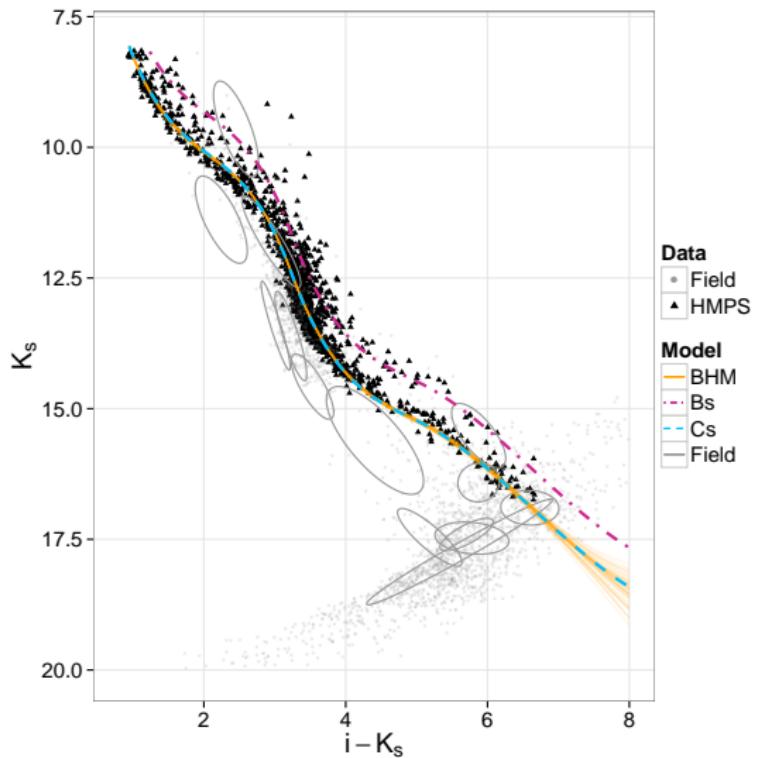
Results



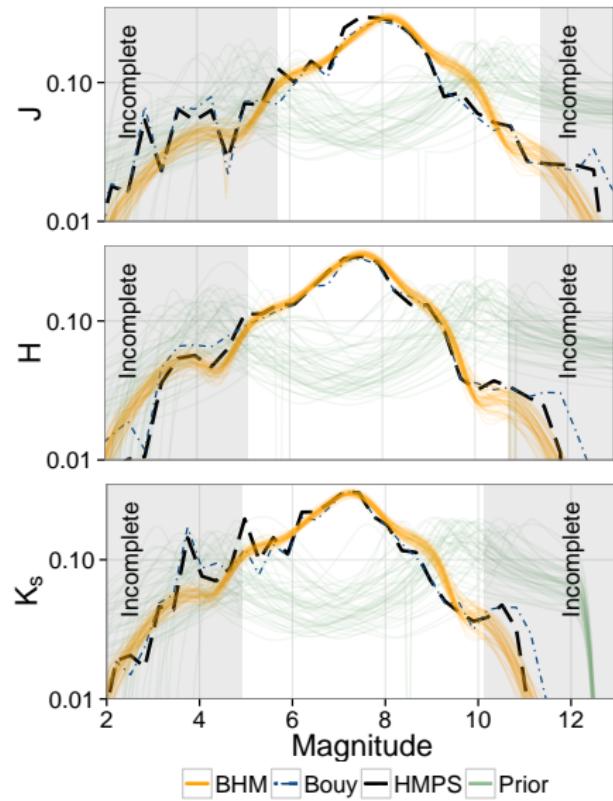
Results



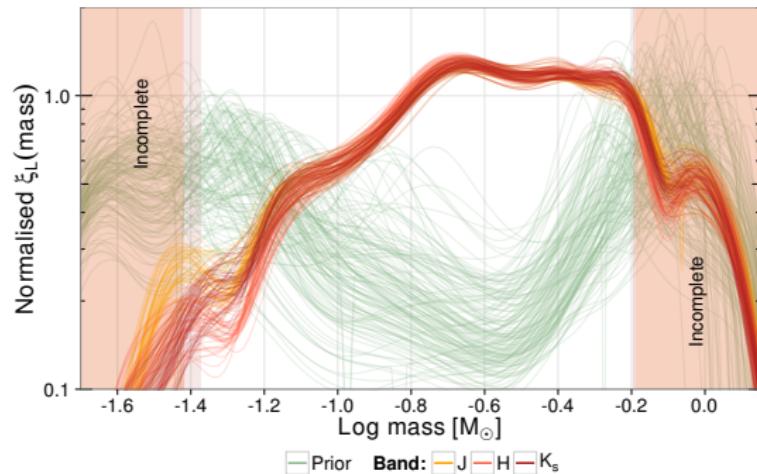
Results



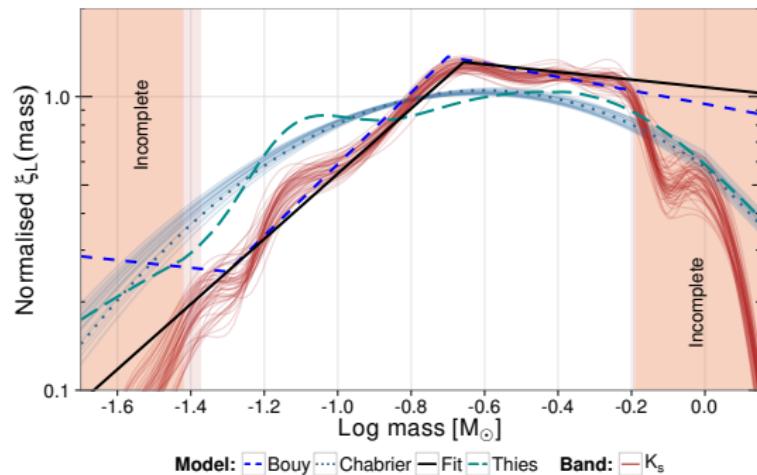
Results



Results



Results





Conclusions



UNED

- Accurate statistical distributions
- Membership probabilities for 10^5 objects
- 205 new members
- Mass distribution with uncertainties from all observables.